A System to Filter Unwanted Messages using Summarization and Classification Techniques

[*]Sheetal M. Rajput[1], Nilima P. Patil[2]

[1](PG Student, Department of Computer Engineering, SSBT's COET, Bambhori, NMU, Jalgaon, India)
[2](Assistant Professor, Department of Computer Engineering, SSBT's COET, Bambhori NMU, Jalgaon, India)
Corresponding Author: Sheetal M. Rajput

Abstract : *Social Network is one of the most popular interactive medium to share, communicate and distribute information. Social Network is the platform to build social relations among people. Users have the ability to keep in touch with friends by exchanging different types of information or messages. Sometimes people post messages which cause a serious problem like harassing or blackmailing to users. The message content may be impolite. The words like offensive, hate, vulgar etc are available in the message. Those messages are detect as spam using information filtering. The Summarization and Machine learning text classification techniques are used for information filtering. The summary of message generated then classification is applied using List Based classification and Neural text classifier. On the basis of text classification checked that whether the messages is spam or spam free.*

Keywords: *Back Propagation Neural Network, Inverse Document Frequency, Neural Text Classifier, Support Vector Machine , Term Frequency*

I. Introduction

Social Networks are today one of the most popular interactive medium to communicate, share, and disseminate a considerable amount of human life information. User has the ability to keep in touch with friends by exchanging different types of information or messages. Sometimes people are post messages which may cause a serious problem like harassing or blackmailing to users [1]. The message content may be impolite. The words like offensive, hate, vulgar etc are available in the message. The messages are detect as spam using information filtering. Information filtering can therefore be used to give users the ability to automatically control the messages written, by filtering out unwanted messages. Text Classification involves assigning a text document to a set of predefined classes automatically, using a machine learning technique. The classification is usually done on the basis of significant words or features extracted from the text document. Since the classes are predefined it is a supervised machine learning task. Most of the official communication and documentation maintained in commercial and governmental organizations is in the form of textual electronic documents and emails. Much of the personal and other communication done by private individuals is in the form of emails, blogs etc. Due to this information overload, efficient classification and retrieval of relevant content has gained significant importance [3]. Text Summarization is condensing the source text into a shorter version preserving its information content and overall meaning. It is very difficult for human beings to manually summarize large documents of text. Text Summarization methods can be classified into extractive and abstractive summarization. An extractive summarization method consists of selecting important sentences, paragraphs etc. from the original document and concatenating them into shorter form. The importance of sentences is decided based on statistical and linguistic features of sentences [6]. An abstractive summarization method consists of understanding the original text and retelling it in fewer words. It uses linguistic methods to examine and interpret the text and then to find the new concepts and expressions to best describe it by generating a new shorter text that conveys the most important information from the original text document.

In social network there is the problem of unwanted messages. The unwanted messages have words like offensive, hate, vulgar etc. To avoid this problem, the paper present the proposed system which able to filter unwanted messages. The spam detection is done by Classification Techniques. The messages have lots of information which are not related to message. Because of such information the number of lines may be increases. To abstract the message, the Summarization Technique is used. There is not more work done on applying the Summarization technique with Text Classification for unwanted message filtering. The summary generated using Summarization Technique and then Classification is done for spam detection. The Repetition and Semantic Matching Techniques are applied for summarization of text input. The classification techniques are applied on summarized input text. The List Based Classification and Neural Text Classifier are used for the spam detection.

II. Literature Survey

Dalal et al., in [4], described Text classification. It refers to the process of assign a category or some categories among predefined ones to each document, automatically. Text categorization is a pattern classification task for text mining and necessary for efficient management of textual information systems. There are two types of approaches to text categorization: rule based and machine learning based approaches. Rule based approaches mean ones where classification rules are defined manually in the form of if then else, and documents are classified based on the rules. This class of approaches has high precision but poor recall, because of its poor flexibility. Machine learning based approaches mean ones where classification rules or equations are defined automatically using sample labeled documents. This class of approaches has a much higher recall but a slightly lower precision than rule based approaches. In addition to their poor flexibility, rule based approaches require time consuming manual jobs for building classification rules. Therefore, machine learning based approaches are replacing rule based ones for text categorization.

M. Ramya et al., in [7], analyze the text summarization techniques. The process of summarization condenses a source document into meaningful content which reflects main thought in the document without altering information. Thus it helps user to grab the main information within short time span. If the user gets effective summary it helps to understand document at a glance without checking it entirely, so time and efforts could be saved. Text summarization process works in three steps analysis, transformation and synthesis. Analysis step analyzes source text and select attributes. Transformation step transforms the result of analysis and finally representation of summary is done in synthesis step. Text summarization approaches generally categorized into: extractive summarization and abstractive summarization. Extractive summarizations extract important sentences or phrases from the source documents and group them to generate summary without changing the source text. However, abstractive summarization consists of understanding the source text by using the linguistic method to interpret and examine the text [8].

III. Proposed Solution

3.1 Proposed Method

The Neural Text Classifier (NTC) describes in Figure 3.1, with respect to its architecture. The proposed neural network follows Perceptron in that synaptic weights are connected directly between the input layer and the output layer, and the weights are updated only when each training example is misclassified. The architecture of the neural network (NTC). It consists of the three layers: input layer, output layer, and learning layer. The input layer receives an input vector given as a string vector. The learning layer determines weights between the input and the output layer corresponding to words of the given input vector by looking up in the tables owned by learning nodes. The output layer generates the categorical scores indicating memberships of the string vector in categories as the output [9].
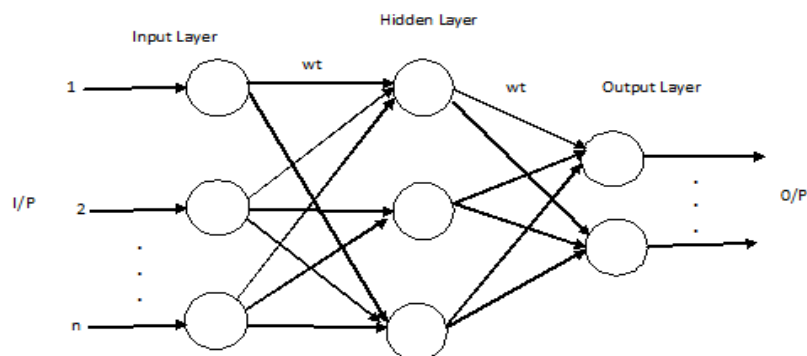

Figure 3.1: Architecture of Neural Text Classifier

3.2 Architecture

A System to Filtering Unwanted Messages using Summarization and Classification Techniques is shown in Figure 2. The input text is given to the system; firstly the words are separated from each sentence. The words are called as Tokens. The Summarization Technique is A System to Filtering Unwanted Messages using Summarization and Classification Techniques applied on it. The summary generated by using Repetition and Semantic Matching. The Classification is apply on Summary to detect the given post is Spam or Spam free. The two Classification Techniques are used for Spam detection. The List Based Classification and Neural Text Classifier get the summarized text as input.
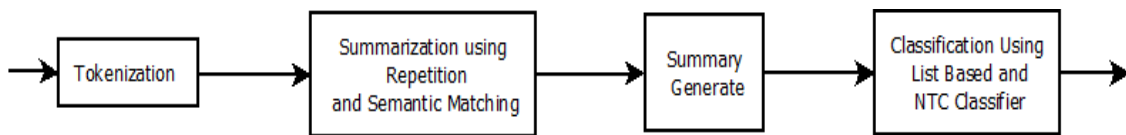
Figure 3.2: Architecture of Proposed System

The Repetition and Semantic Matching used for sentence cohesion. Each word of sentence is matched with word of next sentence. If the semantically matching is there in two sentences then add into summary. In this way the summary is generated. For example Sent A, Sent B, Sent C, Sent D are the sentences. The Repetition technique is applied on sentences like comparing the similarities of Sent A and Sent B, so on...

Now the next process is classification. In the classification process, the summarized text is given to classification techniques. Here two classification techniques are used. First is List Based Technique. The List based classification is checked that whether the word is unwanted or not. If the word of sentence is offensive, hate, vulgar etc. then detect as it is spam.

The Neural Text Classifier is also used for classification. In Neural Text Classification process the training is apply on the summarized data. In the training process, the stop words are removed from given input. Common words with no semantics and which do not provide important information for the final summary are eliminated. The Back Propagation Neural Network (BPNN) algorithm is used to classify the text as spam or not spam.

The neuron present in hidden layer. Weights are given to nodes.
Z_j = hidden unit j. The net input to Z_j is

$$Z_{inj} = V_{oj} + \sum X_i; V_{ij} \tag{3.1}$$

Where, X = input training vector
$\qquad$ t = target output vector
$\qquad$ Z_j = hidden unit j
Y_k is output unit k. The net output to Y_k is

$$Y_{ink} = W_{0k} + \sum Z_j W_{jk} \tag{3.2}$$

3.3 Design

The text input is given to the proposed system. The tokens are created then the tokens are comparing with words of next sentence to find the repetition. According to the condition the sentences are added into the summary. The summarized text is then classified by List Based Classification and Neural Text Classifier. On the basis of classification, the output message is displayed as spam or spam free.

3.3.1 Algorithm for Summarization
The algorithm for summarization is shown. The count of repetition of words in one sentence to the next sentence is calculated. The count value is compared with threshold value. According to the given condition it is add into summary.
1: Input: The sentences or paragraph.
2: Output: Summary of given input.
3: Create the tokens for each sentence and store in vector.
4: Remove Stop Words like: the, and, or, but etc.
5: Compare each word of one sentence to the words of another sentence.
6: If Word [A] is equal to Word [B] then Sentence A and Sentence B are add into the summary.
7: Else only Sentence A adds into summary.

3.3.2 Algorithm for List Based Classification
The algorithm for List Based Classification is as Algorithm 2. In this classification the list of offensive, hate, vulgar etc. words are provided. The words of each sentence of summarized input are checked with these unwanted words. If such words are found then display the message as Spam or Spam free.
1: Input: Generated Summary is input for classification and the files of words offensive, hate, vulgar.
2: Output: Spam detects.
3: Create tokens using blank spaces.
4: Remove Stop Words like: is, the, and, or, but etc.
5: Read files which have list of unwanted words.

6: Check the each word of summary with words of given list.
7: If any word is matched with the word of list.
8: Display message as Spam with its category.
9: Else display message as spam free.

4   Results and Discussion

4.1 Performance Metrics

Performance Metric used in filtering unwanted messages (Spam detection) are Precision, Recall and finally the F-Measure from respective precision and recall values.

Precision =No. of relevant records retrieved / (No. of relevant records retrieved + No. of irrelevant records retrieved)

Recall =No. of relevant records retrieved / (No. of relevant records retrieved + No. of relevant records not retrieved)

F-Measure = (2 * Precision * Recall) / (Precision + Recall)

4.2 Experimental Results

Table 4.1: Result Analysis for Neural Text Classifier

Neural Text Classifier	Precision	Recall	F-Measure
Input 1	85	78	81.34
Input 2	82	74	77.79
Input 3	80	72	75.78
Input 4	85	80	82.42
Input 5	84	78	80.88

The results analysis for Neural Text Classifier is shown in Table 4.1. The five different inputs are given to NTC classifier and get the results. The values of Precision, Recall and F-measure are determined. The average Precision value is 83.2 for NTC, Recall value is 76.4 for NTC. The accuracy of NTC is 80%. The results of Precision, Recall and F-Measure are graphically represented as follows:
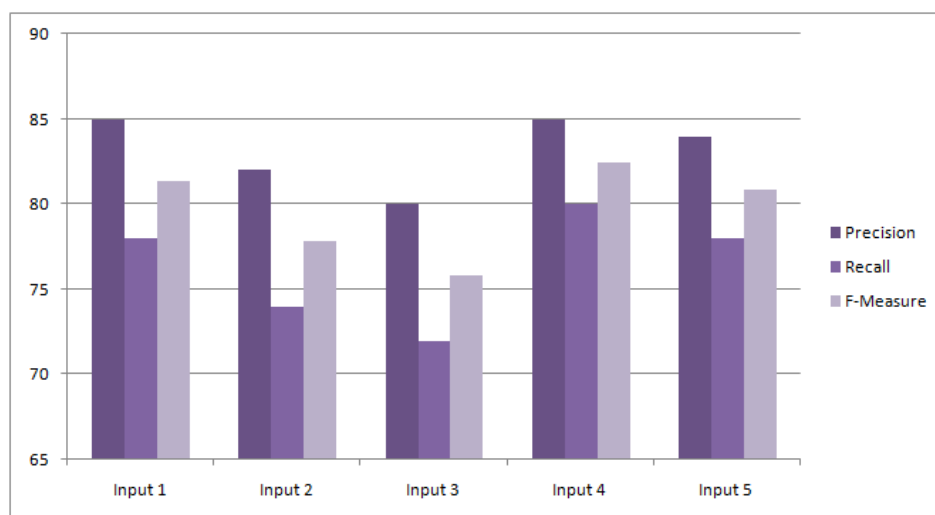

Figure 4.1: Precision, Recall and F-measure for NTC

4.3 Discussion

The Spam detection system is developed that uses the summarization and two classification techniques. The summary is generated from given input text. The summarization uses Repetition and Semantic matching technique. The summary is given to List Based Classification and Neural text classifier. The list of different unwanted words is used for List Based Classification. The Neural Text Classifier used the Back Propagation algorithm for classification. The system determines whether the post is SPAM or SPAM Free. The Precision and Recall is determined by number of relevant records retrieved and number of irrelevant records retrieved. From

results it is found that the accuracy of NTC is better than List Based classification. An accuracy of List Based is 60% and accuracy of NTC is 80%.

5   Conclusion

In this paper, the unwanted messages are filtered by using Summarization and Classification techniques. The important sentences are selected based on Sentence to Sentence Cohesion feature of Extractive Text Summarization using Repetition and Semantic Matching techniques. The generated summary is classified for Spam detection using List Based Classification and Neural Text Classifier. In the summary, the number of sentences are reduced because of that it is beneficial for spam detection. The accuracy of Neural Text Classifier is more than List Based Classification. In Future, the system is implemented by Summarization using Linguistics Cohesion technique.

Acknowledgements

References

[1]. M. Vanetti, E. Binaghi, E. Ferrari, B. Carminati, and M. Carullo, "A system to filter unwanted messages from OSN user walls,"IEEE Transactions on Knowledge and data Engineering, vol. 25, no. 2, pp.285-297, 2013.

[2]. P. Salunkhe, S. Bharne, and P. Padiya, "Filtering unwanted messages from OSN walls," in 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH) IEEE, pp. 261-264.

[3]. P.-y. Zhang and C.-h. Li, "Automatic text summarization based on sentences clustering and extraction," IEEE International Conference on Computer Science and Information Technology. IEEE, 2009, pp. 167-170.

[4]. M. K. Dalal and M. A. Zaveri, "Automatic text classification: a technical review," International Journal of Computer Applications, vol. 28, no. 2, pp. 37-40, 2011.

[5]. A. R. Pal and D. Saha, "An approach to automatic text summarization using wordnet," Advance Computing Conference (IACC), 2014 IEEE International. IEEE, 2014, pp. 1169-1173.

[6]. N. Yadav and N. Chatterjee, "Text summarization using sentiment analysis for DUC data," International Conference on Information Technology (ICIT). IEEE, 2016, pp. 229-234.

[7]. M. Ramya and A. Pinakas, "Different type of feature selection for text classification," IJCTT, vol. 10, no. 2, pp. 102-107, 2014.

[8]. R. Barzilay, K. McKeown, and M. Elhadad, "Sentence fusion for multidocument news summarization," Computational Linguistics, vol. 17, pp. 35-55, 2002

[9]. R. A. Garc__a-Hern_andez and Y. Ledeneva, "Word sequence models for single text summarization," in Second International Conferences on Advances in Computer-Human Interactions, 2009. ACHI'09. IEEE, 2009, pp. 44-48.